www.qconferences.com
www.qconbeijing.com

QCon北京2014大会 4月25—27日

伦敦 | 北京 | 东京 | 纽约 | 圣保罗 | 上海 | 旧金山

London · Beijing · Tokyo · New York · Sao Paulo · Shanghai · San Francisco

QCon全球软件开发大会

International Software Development Conference

VISIT…

www.infoq.com/cn

InfoQ

@InfoQ

infoqchina

软件
正在改变世界！

InfoQ
Enterprise Software Development Community

特别感谢
QCon上海合作伙伴

百年人寿系统架构演变

百年人寿

李杰

百年人寿 AEON LIFE 创新百年 关爱永恒

内容要点

一、百年人寿的故事

二、百年人寿系统高可用性的摸索

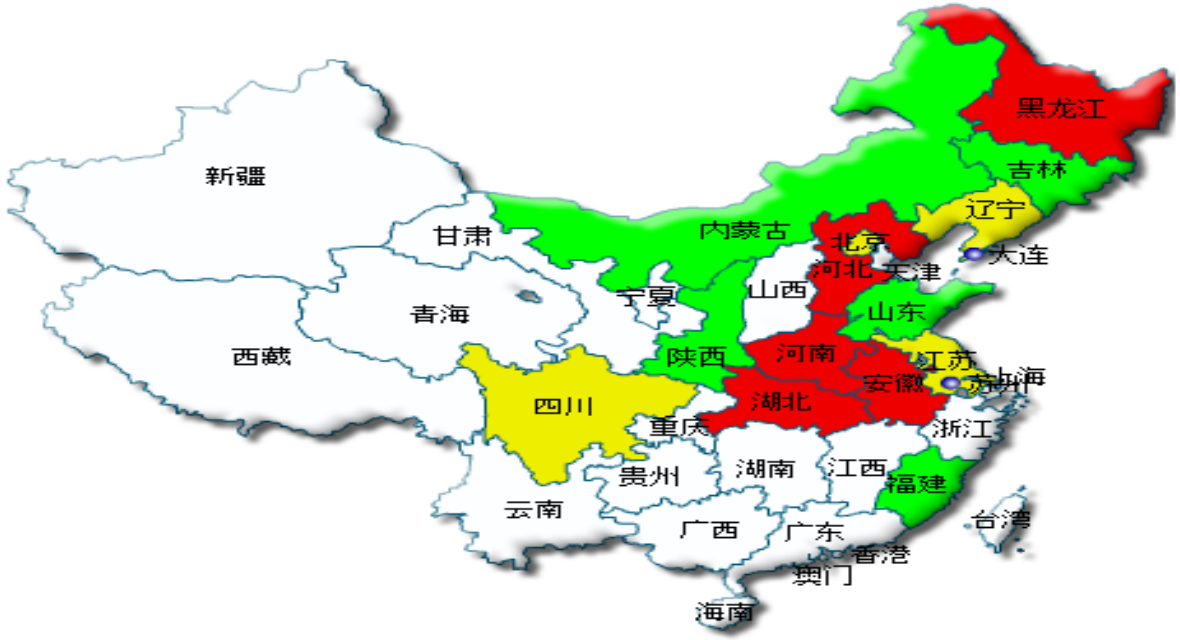
三、未来的展望

百年人寿的故事

2008年批筹，当时公司所有人可在一张桌子上吃饭。

2009年6月开业，2010、2011快速扩张。

2013年机构数量上百家，分布在15个省市。



百年人寿的故事

系统不断增加	核心系统	财务系统	电销系统…
用户不断增多	100	800	3000…
网络不断扩张	大连	湖北	全国…

随着公司业务的不断发展，IT架构也在逐步完善，公司对信息系统和网络的各种要求也越来越高。

我们从最开始就有了规划，伴随着业务的发展来逐步实施

百年人寿的故事

百年人寿系统建设遇到的障碍

快速、持续的开拓新机构，导致系统操作培训等工作跟不上节奏

系统需求及流程不断变化，导致系统更新的相关工作要更加高效

出于成本控制考虑，分支机构IT人员数量少，技能要求相对较低

随着信息化建设的逐步深入，集团型企业在经过“十一五”信息化系统大规模投资建设阶段后，应用系统的集约化不断提高，大集中系统占企业信息化系统的比例也越来越大。

百年人寿的故事

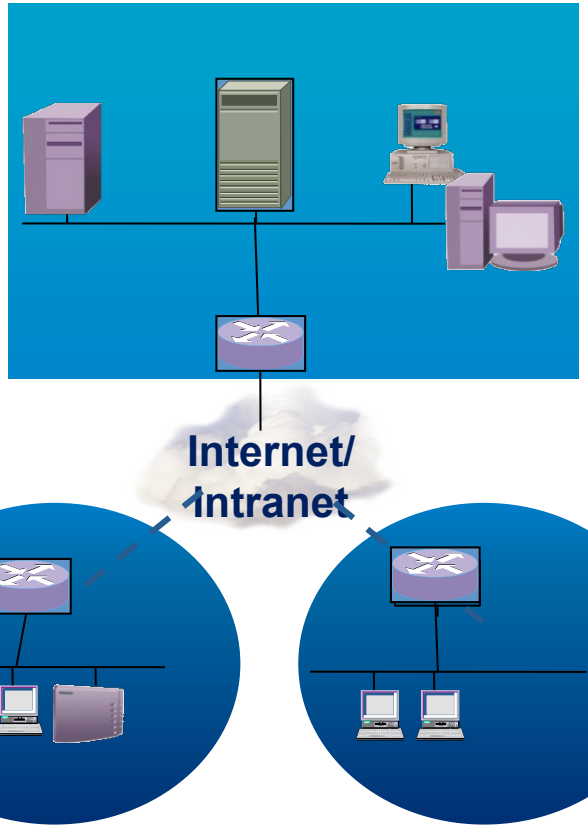
百年人寿系统建设采用完全集中模式

维度	业务管控程度		IT管控程度		
考虑因素	业务	管理	IT管控	基础设施	技术应用
完全集中	**业务操作完全一致**	**实现管控集中**	**应用系统、数据资源、基础设施三方面统一管理**	**集中部署，统一运维**	**接口数量相对较少，对接管理容易**
前置集中	基本业务操作一致，但存在个性化业务应用	实现管控集中	数据资源统一管理、系统程序统一管理	分散部署、分散运维	接口对接数量较多，对接管理相对复杂
区域集中	业务操作实现区域一致	实现区域业务管控集中	应用系统、数据资源、基础设施三方面区域管理	区域部署，区域运维	接口对接数量较多，对接管理相对复杂
物理集中	业务分散处理	分散管理	数据资源集中管理、基础设施统一部署	集中部署，统一运维	接口对接数量较多，对接管理相对复杂

百年人寿的故事

大集中模式---适合自己的架构就是好架构

总公司信息中心



各个分支机构不需部署
业务系统和数据库，
直接在中央系统上作业

各分公司

内容要点

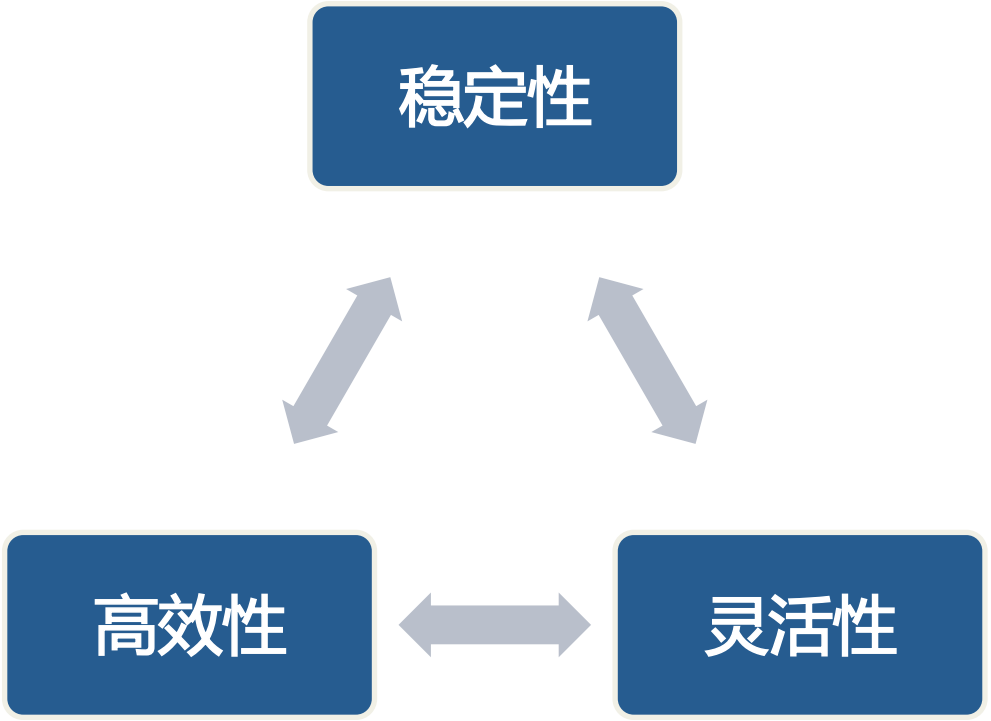
一、百年人寿的故事

二、百年人寿系统高可用性的摸索

三、未来的展望

百年人寿系统高可用性的摸索

客观情况要求系统保持高可用性



百年人寿系统高可用性的摸索

客观情况要求系统保持高可用性

稳定性

公司内部业务操作要求各业务系统保持7X24持续稳定（比如月底冲刺等）

公司对外服务窗口有同样要求（95542电话中心，公司网站，短信平台等）

公司众多内联系统有同样要求（核心、财务、电销、短信等系统互相联系）

公司众多外联系统有同样要求（银行，航意险网站，外包录入，外包打印等）

百年人寿系统高可用性的摸索

客观情况要求系统保持高可用性

高效性

业务量不断增大，历史数据、影像资料需要保留，千万级的数据表不断涌现

系统用户的各种操作、查询和数据分析汇总均要求有非常良好的客户体验

外联系统对核心系统响应时间要求很高（比如银行系统，各种外包服务商等）

外联系统会在短时间内产生大量请求，并要求及时反馈结果（航意险网站等）

百年人寿系统高可用性的摸索

客观情况要求系统保持高可用性

灵活性

公司内部系统较多，各系统之间存在大量的资源共享和互相调用，相关性高

业务需求多变要求相关系统及时作出调整，同时又不能影响周边的其他系统

硬件和网络资源需求在不断增加，要保证相应调整不会影响到用户的使用

系统不论软件架构还是部署方案都需要考虑能够**灵活**适应将来的各种变化

百年人寿系统高可用性的摸索

如何用较低成本满足上述需求，让各个系统顺畅运行，支持用户的日常工作？

百年人寿系统高可用性的摸索

首先看看网络的高可用性

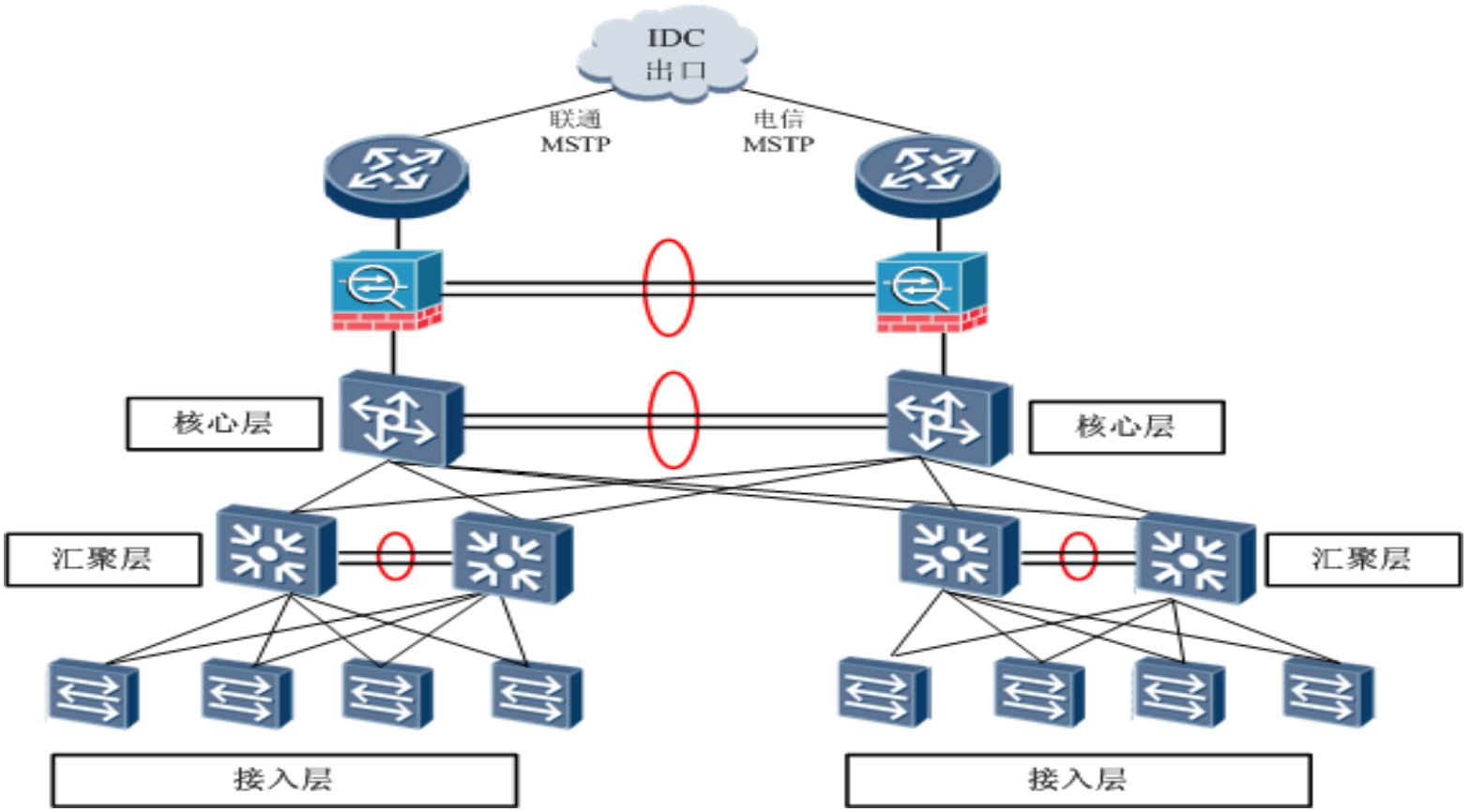
数据中心集中部署，各省分公司及下一级分支机构星形接入，与多家通信运行商合作，通过主备线路和设备冗余配置，实现网络系统稳定高效运行。

采用不少于2条的通信线路，并同时采用不同运营商的线路，避免单运营商、单线路故障。

网络系统设备热冗余配置，多链路设计，路由优化。

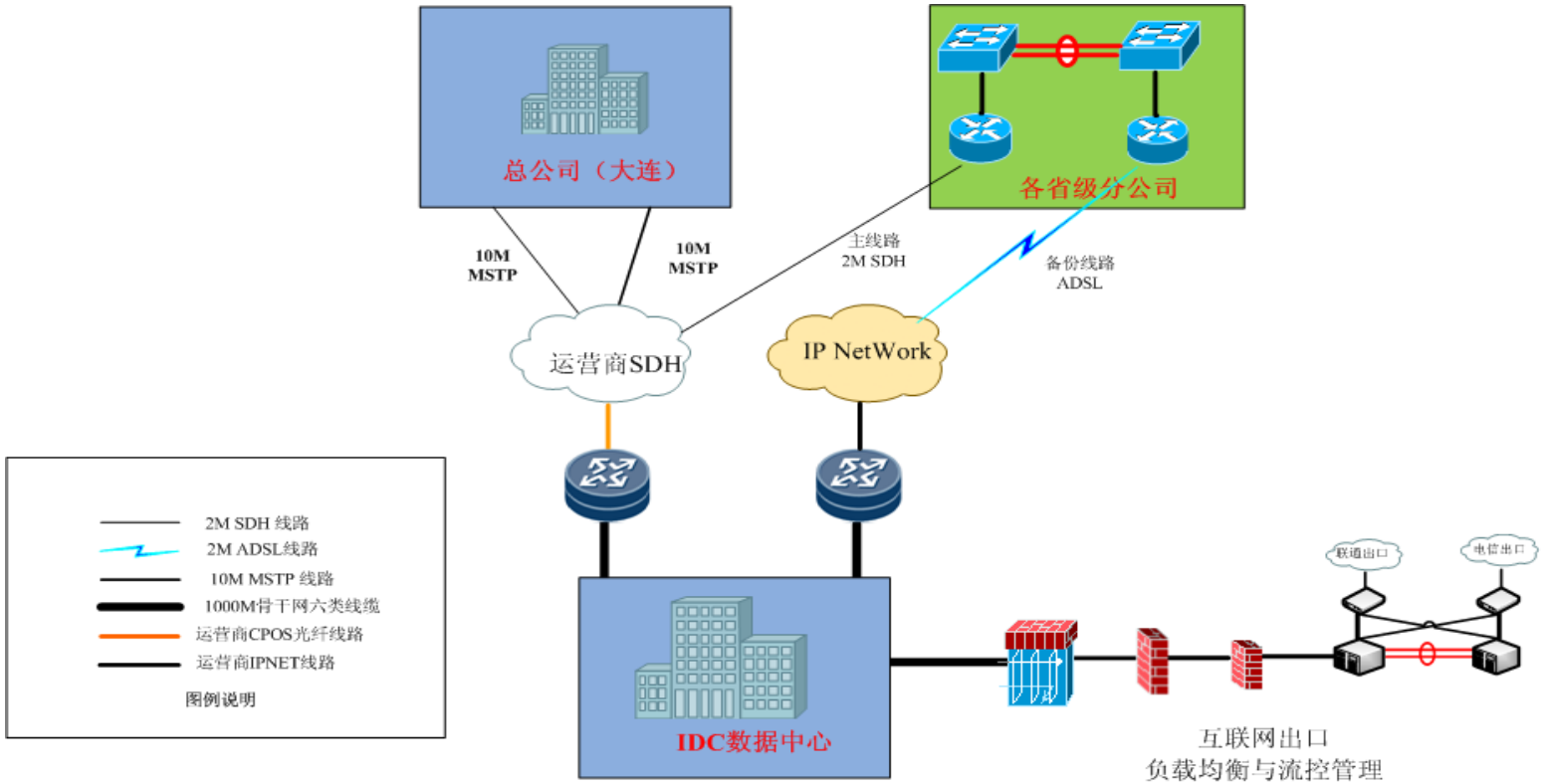
百年人寿系统高可用性的摸索

数据中心网络三层架构



百年人寿系统高可用性的摸索

分支机构网络接入



百年人寿分支机构网络接入拓扑图

百年人寿系统高可用性的摸索

接着看看数据库的高可用性

- 数据库基于HACMP的双机互备
- 基于Goldengate的数据库复制

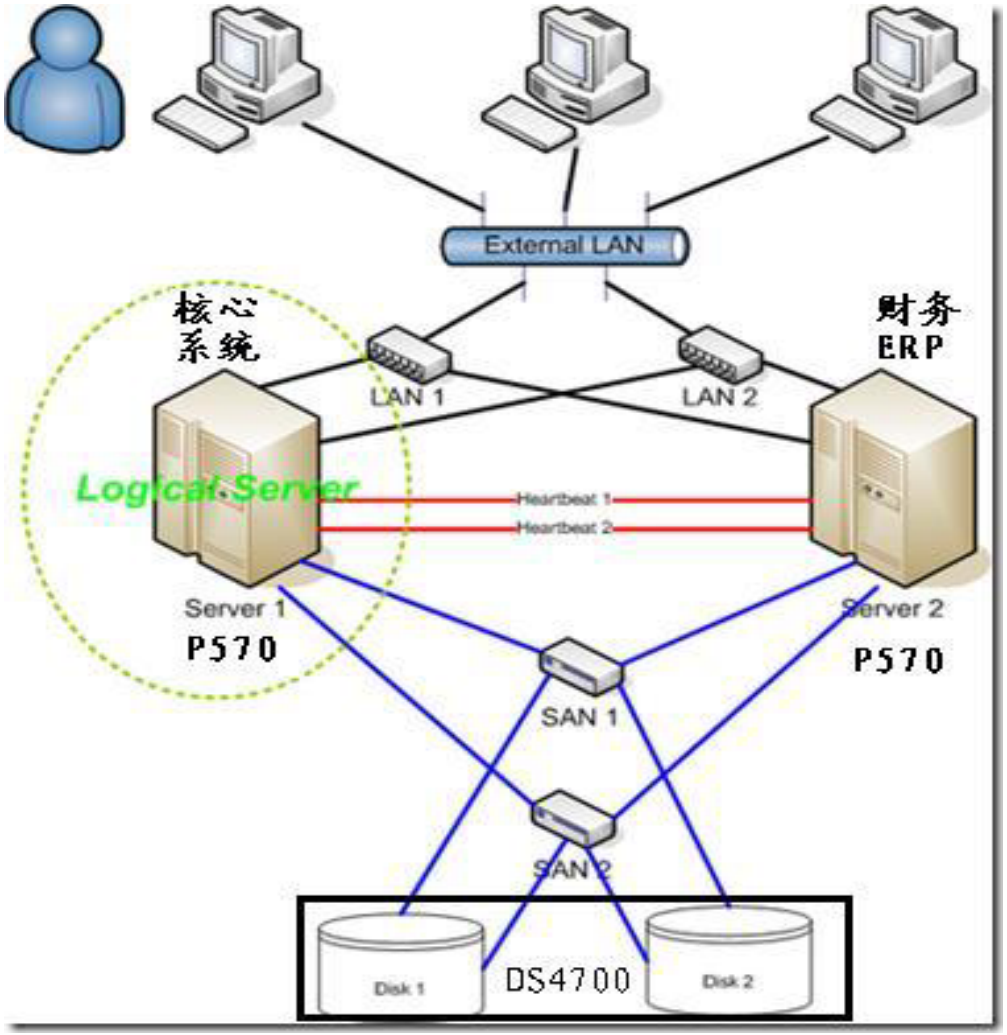
百年人寿系统高可用性的摸索

数据库HA应用情况

1）基于HACMP的双机互备

公司有两大主要系统：**核心业务系统**和**财务ERP系统**，二者的数据库分别运行在2台IBM P570小型机上，采用双机HACMP互备。

两台小型机通过光纤交换机(B24)连接公共磁盘阵列(IBM DS4700)，用该阵列存储数据库的数据文件。



百年人寿系统高可用性的摸索

数据库HA应用情况

在保证高可用性的同时节省成本，充分利用资源

1）基于HACMP的双机互备

两台服务器互为对方的备机，Server1运行核心系统数据库，Server2运行财务ERP系统数据库。当其中一台服务器维护或出现故障时，另一台可以接管相应的数据库服务。比如。当Server2需要宕机维护时，Server1运行核心系统数据库服务的同时，接管并提供财务ERP数据库服务（每个数据库服务仍保持原来的IP不变）；Server2维护结束后，可再接管原来的服务。

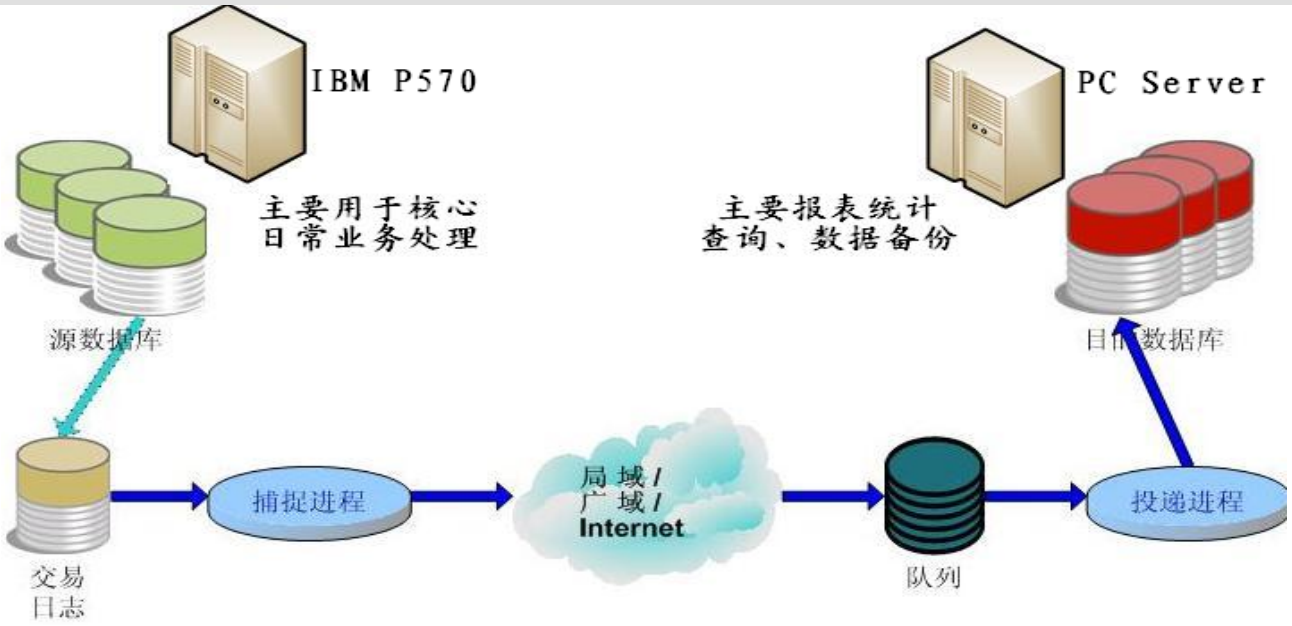
另外，我们也通过RHCS集群来实现数据库的HA，实现原理同HACMP。

百年人寿系统高可用性的摸索

数据库HA应用情况

2）基于Goldengate的数据库复制

我们采用Goldengate对核心数据库实施了同步与复制。源端服务器为P570小型机，目标端为高性能PC Server，实施单向复制。两个服务器部署在同一数据中心，网路延迟很小。日志放在源端，防止网络意外中断。



百年人寿系统高可用性的摸索

最后看看应用部署的高可用性

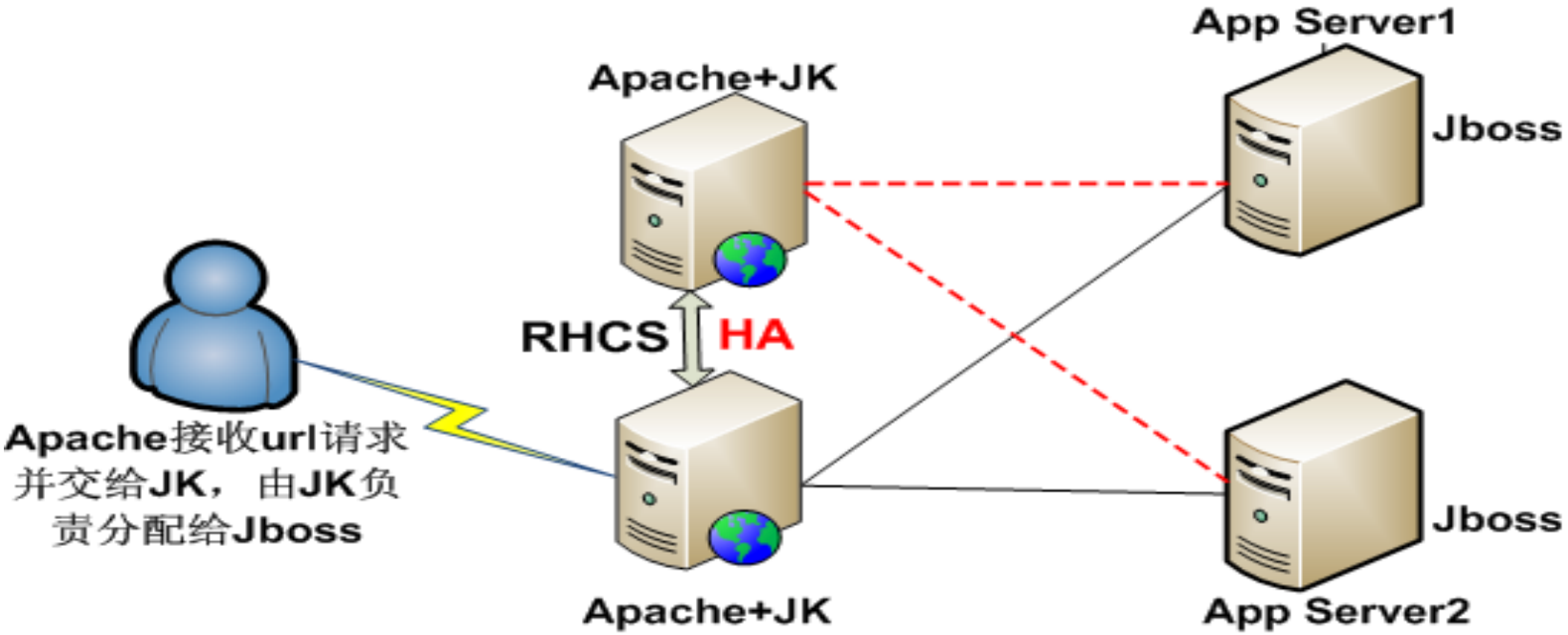
RHCS + Apache+JK实现HA和负载均衡

RHCS+Jboss 实现应用的集群与Session复制

百年人寿系统高可用性的摸索

RHCS+Apache+JK　实现HA和负载均衡

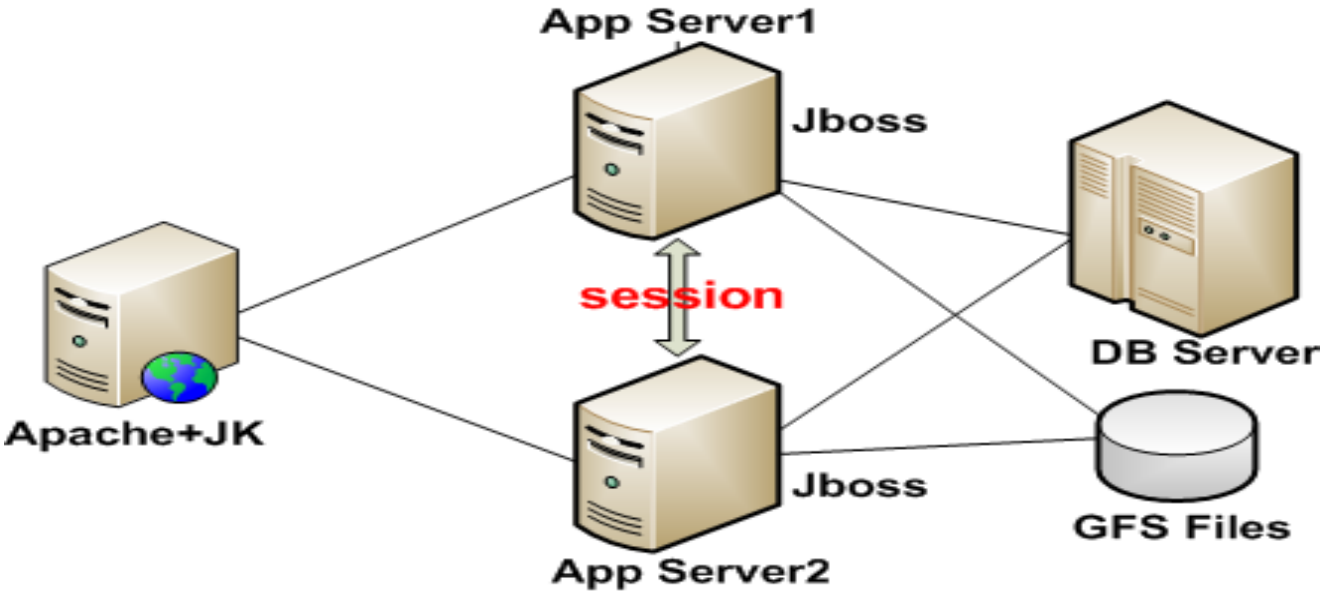
通过RHCS实现服务器集群，其服务资源为服务IP和Apache服务；此架构用以保证负载分发服务器的高可用性。



百年人寿系统高可用性的摸索

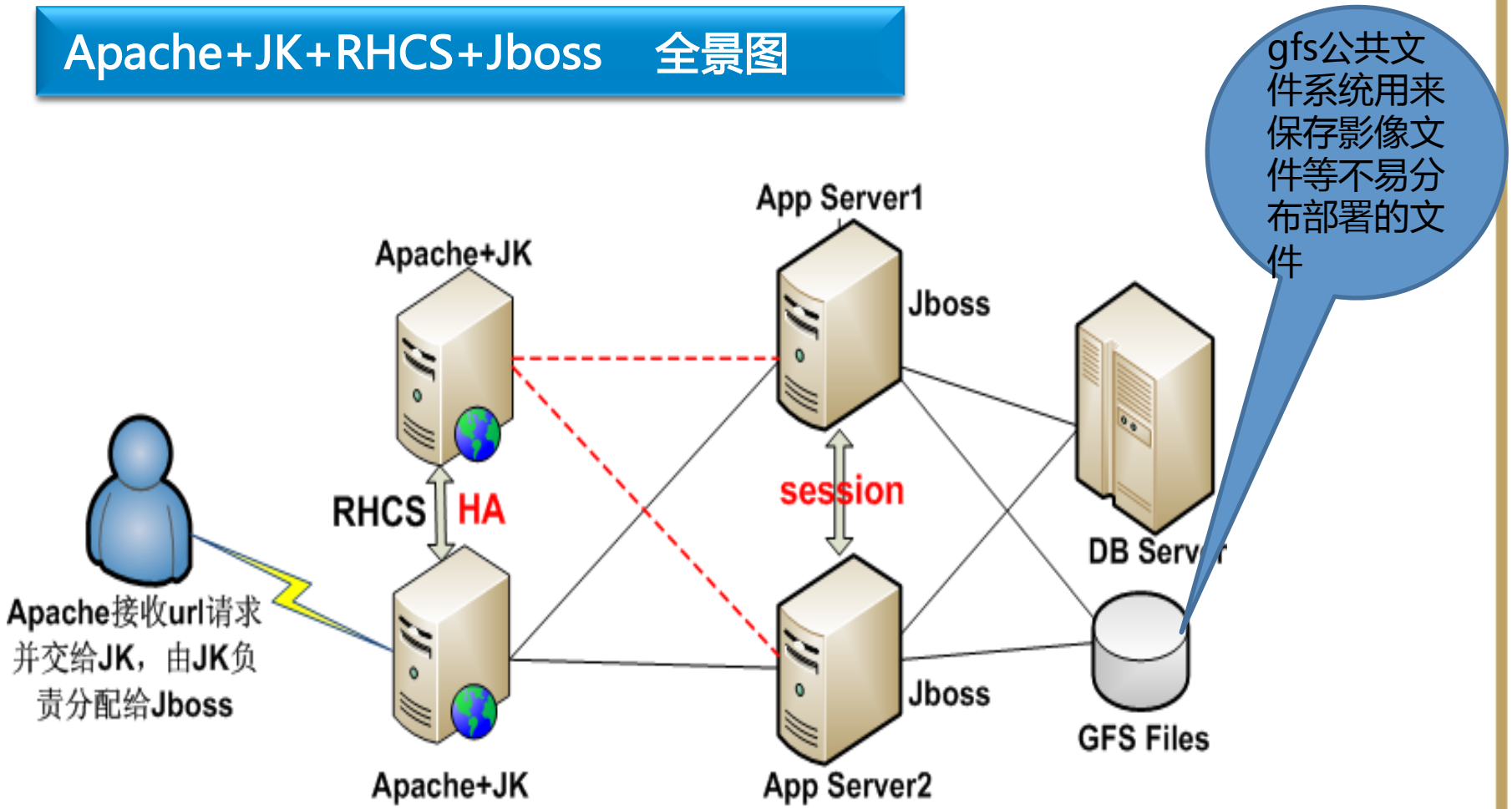
RHCS+Jboss 实现应用的集群与Session复制

Session复制可以避免中断用户访问请求，当一台应用服务器维护或出现故障时，应用的集群和session复制可以保证从一台服务器到另一台的自动切换，而对用户感觉而言，这种切换是“无变化”的。



百年人寿系统高可用性的摸索

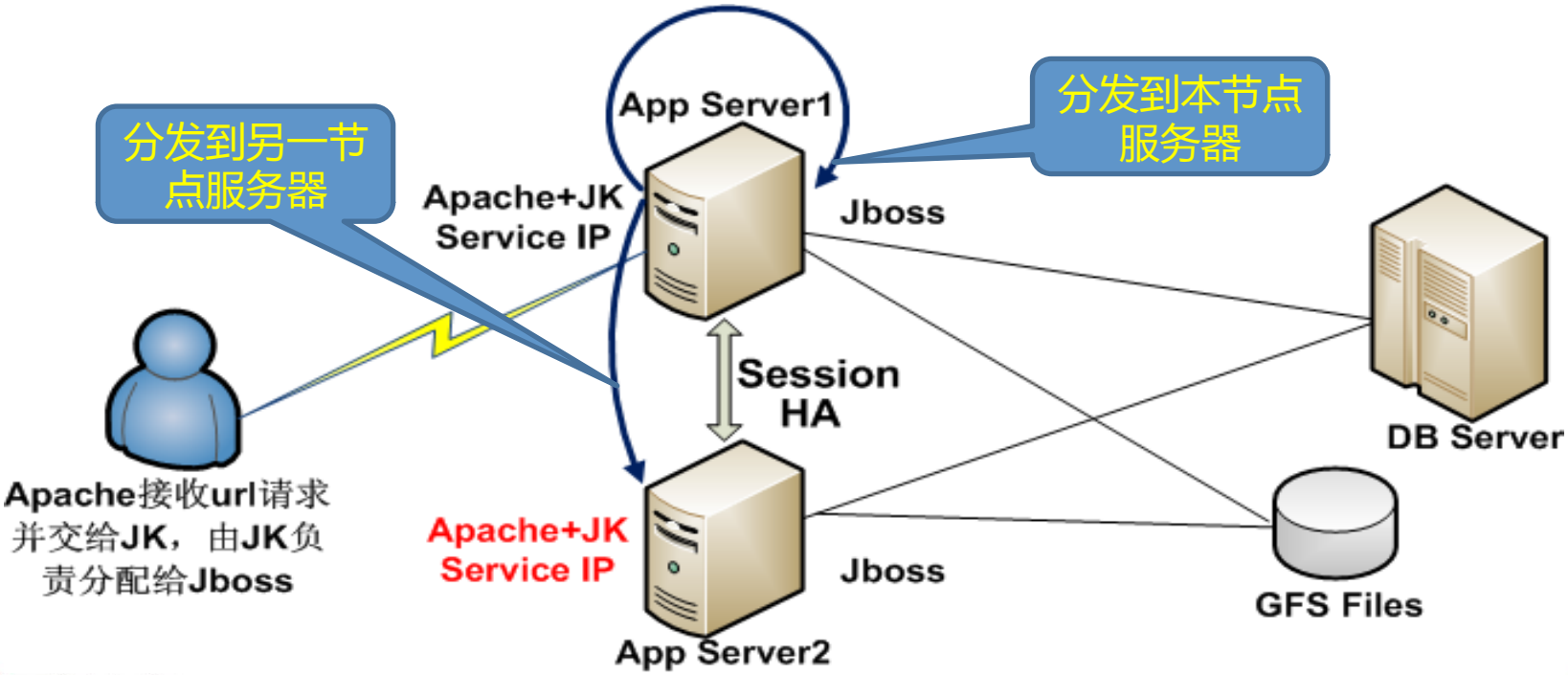
Apache+JK+RHCS+Jboss　全景图



百年人寿系统高可用性的摸索

百年人寿的实际部署情况

每台应用服务器均安装和配置Apache和JK（mod_jk)，可以同时充当应用服务器和负载均衡分发服务器两种角色，结合RHCS的集群服务，即可全面实现系统的HA、负载均衡和应用集群。这样，即使只有一台服务器正常工作，也可以保证应用系统的稳定性和连续性。



百年人寿系统高可用性的摸索

实时监控作为最后的防线

通过Nagios监控，提前预警和解决运行时问题

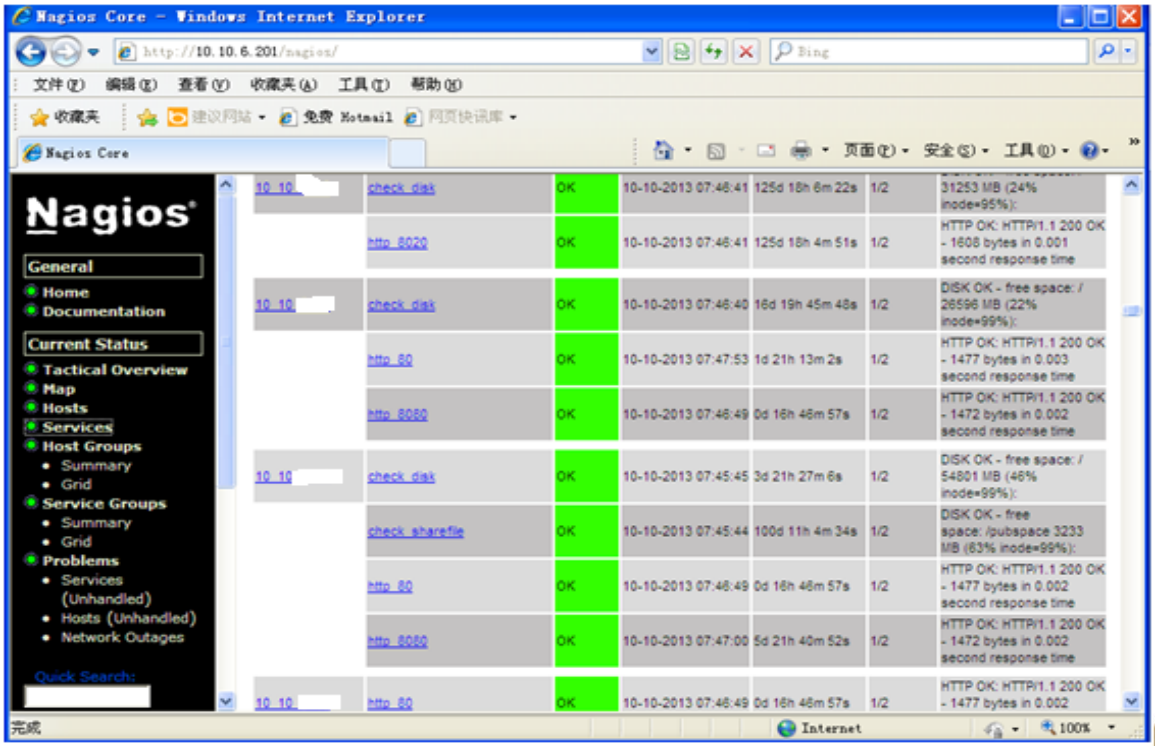
部署Compuware监控系统，捕捉用户使用的实时感受

百年人寿系统高可用性的摸索

Nagios监控

通过开源监控软件（Nagios）实时监控应用服务运行状态、服务器参数指标以及Oracle数据库的相关运行情况。

- 不同服务器可以定制不同的监控指标和预警阀值参数。
- 可配置多个用户，监控和管理不同的服务器或服务。
- 可根据需要将系统管理员进行分组，预警信息通知给需要的管理员和管理员组。
- 可通过邮件和短信发送预警。



百年人寿系统高可用性的摸索

Nagios监控

1）、监控服务器状态

10_10	UP	10-10-2013 08:07:48	447d 23h 10m 32s	PING OK - Packet loss = 0%, RTA = 0.40 ms
10_10	UP	10-10-2013 08:07:48	447d 23h 10m 29s	PING OK - Packet loss = 0%, RTA = 0.67 ms
10_10	UP	10-10-2013 08:08:28	315d 21h 6m 32s	PING OK - Packet loss = 0%, RTA = 0.17 ms
10_10	UP	10-10-2013 08:08:28	315d 21h 5m 46s	PING OK - Packet loss = 0%, RTA = 0.12 ms
10_10	UP	10-10-2013 08:07:48	315d 21h 5m 0s	PING OK - Packet loss = 0%, RTA = 0.19 ms

2）、监控http服务和磁盘空间

10_10	check_disk	OK	10-10-2013 08:10:40	64d 15h 4m 28s	1/2	DISK OK - free space: / 157087 MB (65% inode=99%):
	check_image	OK	10-10-2013 08:10:40	64d 15h 4m 9s	1/2	DISK OK - free space: /lisimage 822147 MB (43% inode=94%):
	http_80	OK	10-10-2013 08:10:40	22d 14h 26m 2s	1/2	HTTP OK: HTTP/1.1 200 OK - 2799 bytes in 2.061 second response time
	http_8080	OK	10-10-2013 08:09:45	4d 19h 7m 19s	1/2	HTTP OK: HTTP/1.1 200 OK - 2794 bytes in 0.005 second response time

百年人寿系统高可用性的摸索

Nagios监控

3）、监控数据库表空间

10_10	Tablespace_Usage	OK	10-10-2013 08:13:07	3d 11h 13m 4s	1/2	OK - tbs USERS usage is 0.00%, tbs UNDOTBS1 usage is 0.32%, tbs TEMP usage is 0.01%, tbs SYSTEM usage is 1.96%, tbs SYSAUX usage is 4.54%, tbs PUBUNDO usage is 0.01%, tbs PUBTEMP usage is 0.00%, tbs LISTEMP usage is 0.00%, tbs LISDATA usage is 81.56%, tbs GGSA usage is 14.12%

4）、监控服务器参数指标

10_10	C:\ Drive Space	OK	10-10-2013 08:16:41	93d 15h 12m 34s	1/2	c: - total: 39.07 Gb - used: 27.99 Gb (72%) - free 11.08 Gb (28%)
	CPU Load	OK	10-10-2013 08:16:40	93d 15h 14m 59s	1/2	CPU Load 9% (5 min average)
	D:\ Drive Space	OK	10-10-2013 08:15:49	0d 17h 34m 15s	1/2	d: - total: 96.90 Gb - used: 51.20 Gb (53%) - free 45.71 Gb (47%)
	Memory Usage	OK	10-10-2013 08:16:40	93d 15h 13m 46s	1/2	Memory usage: total:8137.48 Mb - used: 3898.82 Mb (48%) - free: 4238.66 Mb (52%)
	NSClient++ Version	OK	10-10-2013 08:16:40	93d 15h 13m 10s	1/2	NSClient++ 0.3.9.327 2011-08-16
	Uptime	OK	10-10-2013 08:16:41	93d 15h 12m 34s	1/2	System Uptime - 165 day(s) 18 hour(s) 13 minute(s)

百年人寿系统高可用性的摸索

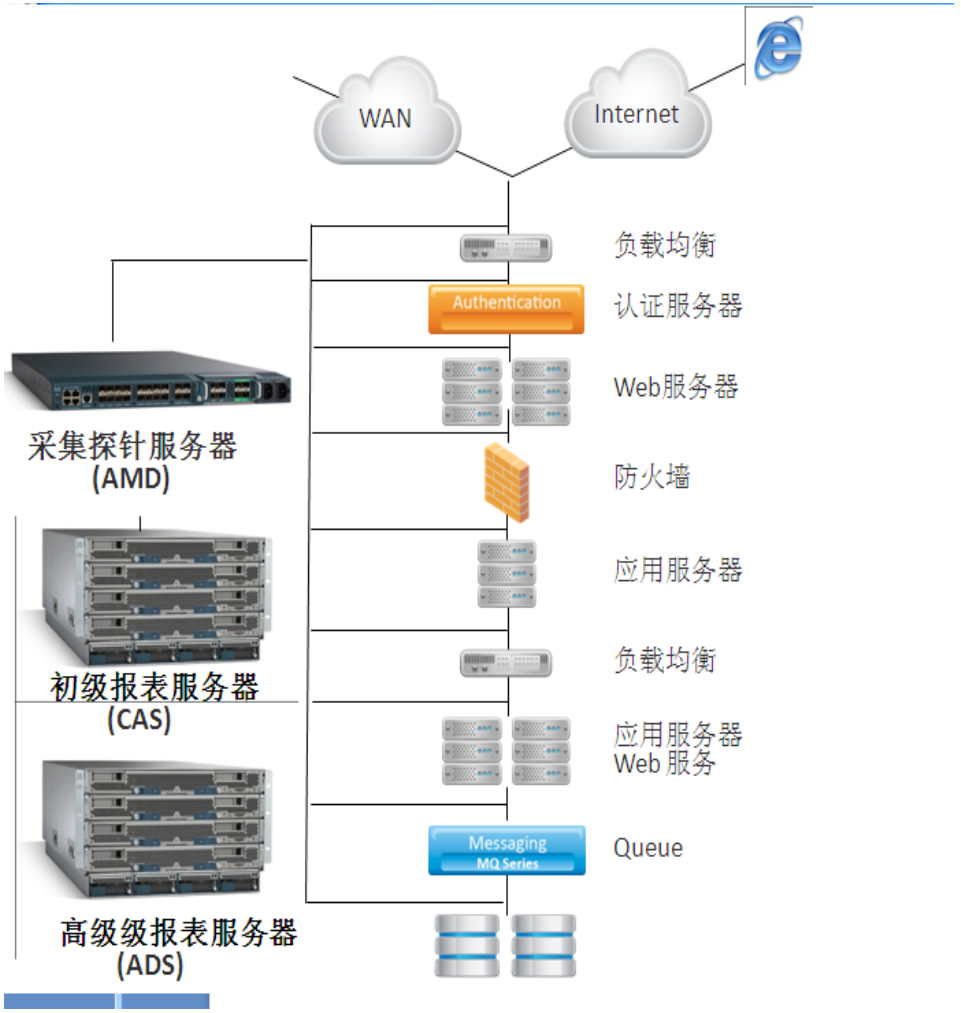
Compuware监控

Compuware监控系统由三台服务器组成，各功能及连接示意图如下：

1）探针服务器连接在网络交换机上，用于数据包抓取

2）初级报表服务器用于分析和统计页面级报表

3）高级报表服务器用于分析和统计元素级报表

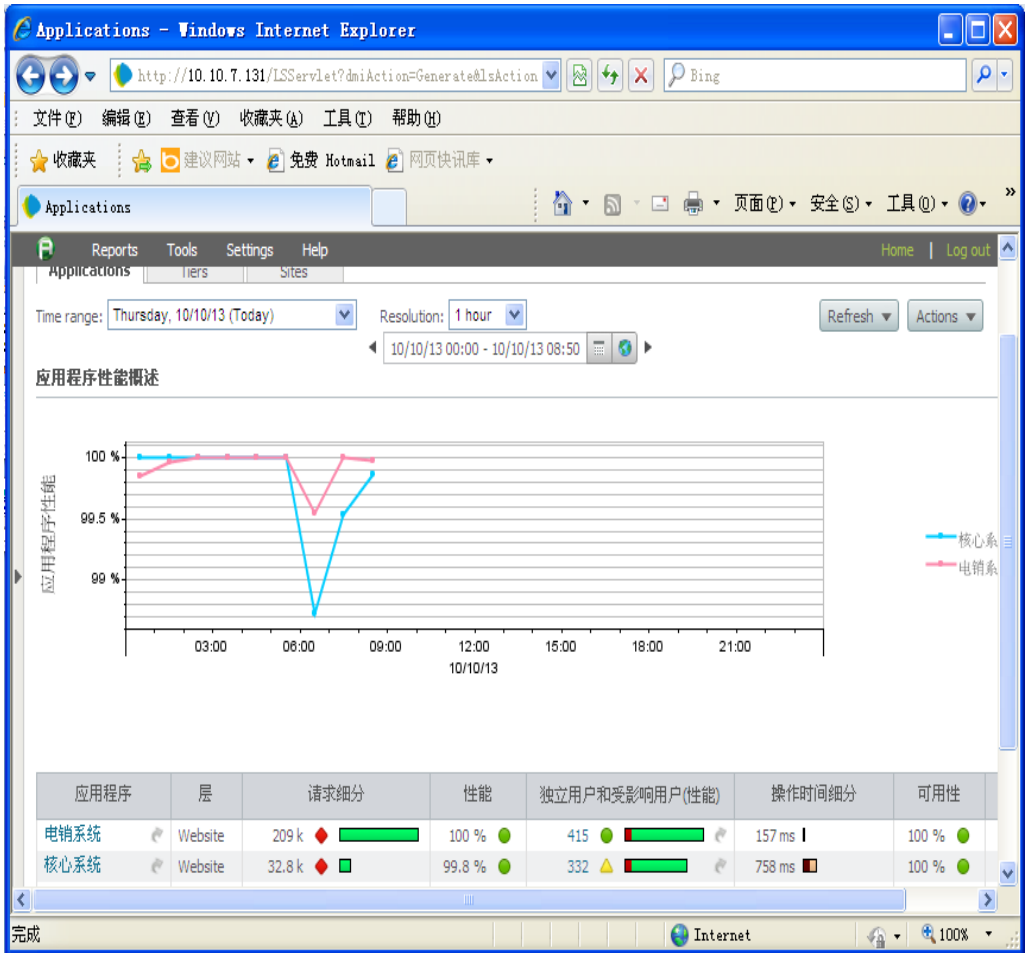


百年人寿系统高可用性的摸索

Compuware监控

部署Compuware监控系统，实时了解用户端PC-网络环境-应用服务器-后台数据库运行，整体链路的各个环节运行状态，及时发现问题并加以解决。

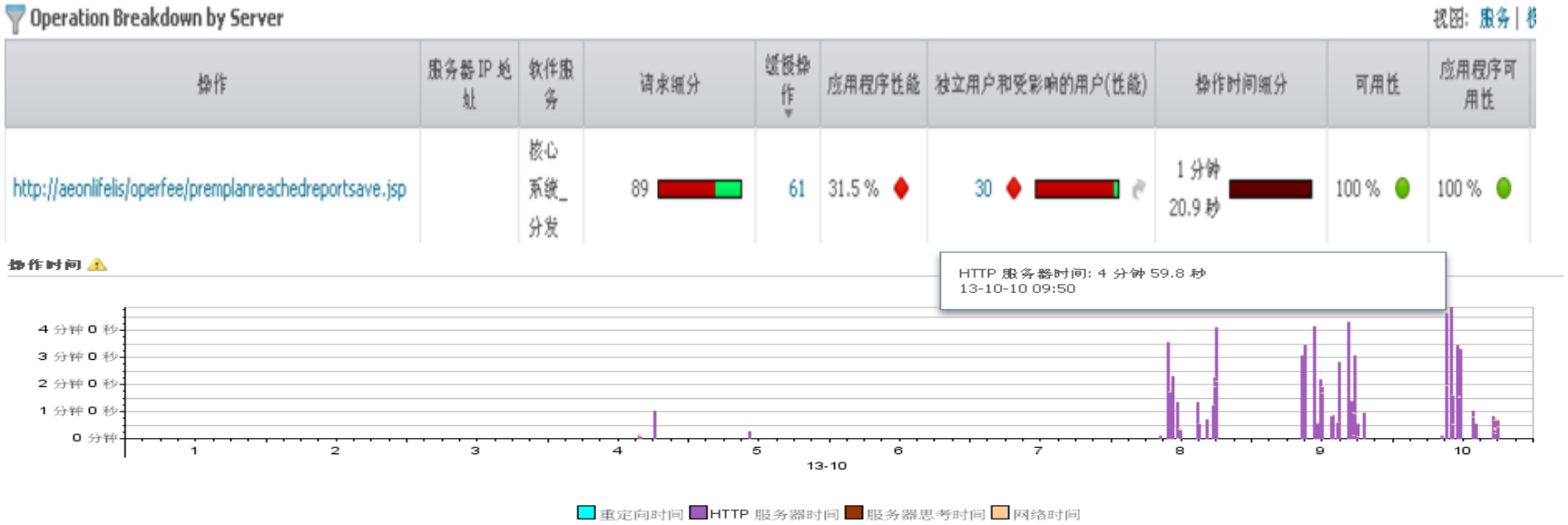
结合我们实际，监控核心业务系统和电销业务系统运行情况，实时监控用户与应用服务器之间的请求、响应的延时情况，监控应用服务器与数据库服务器之间的请求和响应情况，并根据需要设定响应的报警阀值。



百年人寿系统高可用性的摸索

Compuware监控

使用Compuware实时监控和记录用户使用应用系统的操作情况和感受



该页面过去10天有89次请求，其中缓慢操作61次，中止操作0次，受影响用户30个，占比93%，平均操作时间1分钟20.9秒，其中服务器占比100%，网络0%。

内容要点

一、百年人寿的故事

二、百年人寿系统高可用性的摸索

三、未来的展望

百年人寿系统未来的展望

未来规划

- 小型机的更换，增加硬件性能
- 扩展F5应用的范围
- 虚拟技术的应用-VMWare

百年人寿系统高可用性的摸索

小型机的增加

在原有的两台P570的基础上，新增加两台P770。

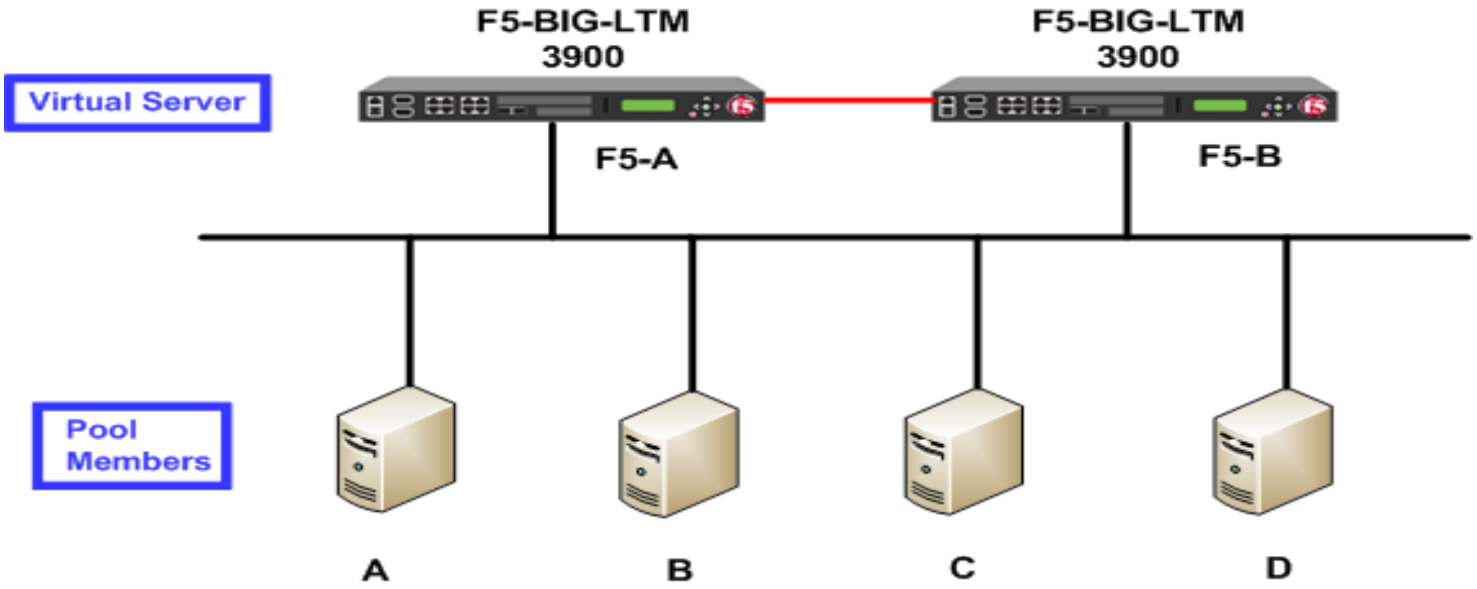
将报表数据库， Goldengate同步数据库等较重要的数据库从普通服务器转为部署在小型机上，使这些数据库能够享有更高的硬件资源和基于HACMP的双机热备服务。

将原先共享HACMP双机热备服务的核心系统数据库和财务ERP数据库拆分开来，与新的更轻量级的其他业务系统的数据库共享HACMP，避免一台小型机出现故障时，另外一台负载过高。

百年人寿系统高可用性的摸索

F5的进一步应用

F5 实现集群与负载均衡
采用两台F5-BIG-LTM-3900实现集群与负载均衡，两台F5设备配置双机热备，以避免F5设备出现单点故障。各应用服务器通过网路分别同时连接两台F5负载均衡设备，物理拓扑图图下：



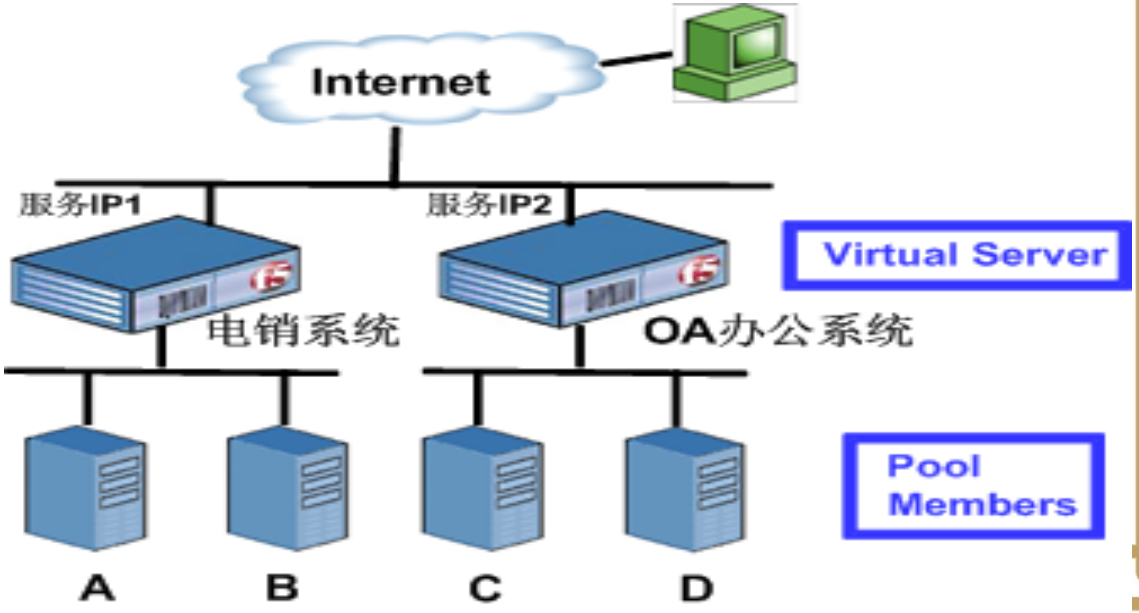
百年人寿系统高可用性的摸索

F5的进一步应用

F5 实现集群与负载均衡
在F5设备中，可以根据实际需要，配置多个Virtual Server（VS），每个VS对应一套应用系统，并拥有一个服务IP，每个VS可以对应多个Members，也就是可以对应多个应用服务器，以完成应用服务器的集群和负载均衡。

我们在实际应用中，已配置了两个VS，一个用于电销系统，另一个用于公司的OA办公系统，目前各有两台应用服务器来完成服务器集群和负载均衡；

后续我们将根据实际需要，增加VS和Member的配置。



百年人寿系统高可用性的摸索

VMWare的应用

充分利用服务器虚拟化功能，建立灵活、易于维护的服务器集群；转变、管理和优化百年的IT基础架构

研究利用桌面虚拟化功能，丰富客户端配置手段，减少运营维护成本。

VMware HA的体系结构和概念提供了基于虚拟化的HA解决方案，将来可能在节约资源的前提下，用于部分或全面替代目前的HA解决方案。

企业云的应用...

结语

负载均衡方面

- F5负载均衡
- Apache+Mod_jk 支持集群和负载均衡

数据库方面

- HA技术
- 数据库同步复制

应用部署方面

- RHCS、Jboss集群
- Sesssion复制

通过集群、HA等技术，从网络、数据库、应用等多方面保障系统的高可用性和稳定性，提高用户满意度。

Q&A？

欢迎探讨，指正！

2013.11

百年人寿
AEON LIFE
创新百年 关爱永恒

InfoQ
Enterprise Software Development Community

特别感谢
QCon上海合作伙伴